

**This Page Is Inserted by IFW Operations  
and is not a part of the Official Record**

## **BEST AVAILABLE IMAGES**

**Defective images within this document are accurate representation of  
The original documents submitted by the applicant.**

**Defects in the images may include (but are not limited to):**

- **BLACK BORDERS**
- **TEXT CUT OFF AT TOP, BOTTOM OR SIDES**
- **FADED TEXT**
- **ILLEGIBLE TEXT**
- **SKEWED/SLANTED IMAGES**
- **COLORLED PHOTOS**
- **BLACK OR VERY BLACK AND WHITE DARK PHOTOS**
- **GRAY SCALE DOCUMENTS**

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**

**THIS PAGE BLANK (USPTO)**

6058 17/00



①9 BUNDESREPUBLIK  
DEUTSCHLAND



DEUTSCHES  
PATENTAMT



Patentschrift  
⑩ DE 197 04 314 C 1

⑥ Int. Cl.<sup>6</sup>  
**G 05 B 13/04**  
// G06F 17/10

②1 Aktenzeichen: 197 04 314.3-51  
②2 Anmeldetag: 5. 2. 97  
④3 Offenlegungstag: -  
④5 Veröffentlichungstag  
der Patenterteilung: 3. 9. 98

(1)

Innerhalb von 3 Monaten nach Veröffentlichung der Erteilung kann Einspruch erhoben werden

<p>⑦3 Patentinhaber: Siemens AG, 80333 München, DE</p>	<p>⑦2 Erfinder: Weinzierl, Klaus, Dr., 81671 München, DE; Frankl, Astrid, 80634 München, DE</p> <p>⑤6 Für die Beurteilung der Patentfähigkeit in Betracht gezogene Druckschriften: WO 96 28 771 Unbehauen, R., "Systemtheorie", R. Oldenburg Verlag, München, Wien, 1990, S. 530-545;</p>
<p>⑤4 Verfahren zur Optimierung von Prozeßparametern auf der Basis einer Gradientenschätzung auf einem Rechner</p> <p>⑤7 Das Verfahren nutzt die Gradienteninformation zur Minimierung gestörter Zielfunktionen. Dabei wird der Gradient ohne zusätzliche Zielfunktionsauswertungen mittels RLS-Algorithmus geschätzt und mit einem stochastischen Minimierungsverfahren kombiniert. Anwendung des Verfahrens ist die Optimierung von Modellparametern eines Kühlstreckenmodells zur Minimierung des Modellfehlers.</p>	

DE 197 04 314 C 1

DE 197 04 314 C 1

In der Industrie werden häufig Prozeßmodelle zur Beschreibung technischer Anlagen eingesetzt. Dabei sind noch unbekannte Prozeßparameter zu bestimmen und die Werte ungenau bekannter Parameter sind zu verbessern.

5 Zur Lösung überbestimmter Gleichungssysteme, die keine exakte Lösung besitzen, kann vorteilhaft mittels globaler stochastischer Optimierungsverfahren eine geeignete Lösung numerisch durch einen Rechner ermittelt werden. Dies ist vor allem bei Prozeßmodellen sinnvoll, die eine große Anzahl Parameter besitzen. Standardalgorithmen zur Minimierung des Modellfehlers scheitern an zahlreichen lokalen Minima, an numerischen Problemen bei der Gradientenberechnung oder an der Rechenzeit.

10 Die Aufgabe des erfindungsgemäßen Verfahrens besteht darin, Prozeßparameter numerisch zu optimieren.

Diese Aufgabe wird durch das Verfahren des Patentanspruchs 1 gelöst.

Zur Ermittlung oder Verbesserung dieser Prozeßparameter werden an der technische Anlage zahlreiche Meßdaten aufgenommen und die Ausgangsgrößen der Anlage mit den Modellausgaben verglichen. Durch Minimierung des Modellfehlers erhält man die gesuchten Parameterwerte.

15 Das erfindungsgemäße Verfahren kombiniert ein stochastisches Optimierungsverfahren mit einem in der Regelungstechnik und Nachrichtenübertragung bekannten Schätzalgorithmus, der zur Schätzung einer Abstiegsrichtung (negativer Gradient) umformuliert wurde. Das stochastische Optimierungsverfahren setzt dabei voraus, daß die Zielfunktion zweimal stetig differenzierbar ist.

Da in der Praxis diese Voraussetzung an die Zielfunktion selten erfüllt ist, ist die numerische Gradientenberechnung oft problematisch. Dies trifft vor allem auf Optimierungsprobleme zu, bei denen die Zielfunktion beispielsweise durch das Softwarepaket des Kunden spezifiziert ist. Bei der dem erfindungsgemäßen Verfahren zugrunde liegenden Minimierung genügt die Kenntnis eines Gradienten eines zweimal stetig differenzierbaren Anteils. Dieser Gradient kann mit dem RLS-Verfahren robuster als mit vergleichbaren Methoden der numerischen Mathematik geschätzt werden. Dabei ist es vorteilhaft, daß der Gradient ausschließlich aus zurückliegenden Funktionsaufrufen ermittelt wird.

25 Es ist vorteilhaft, wenn der Schätzalgorithmus zur Schätzung der Abstiegsrichtung mit einer Zufallssuche kombiniert wird, da so die Suche nach dem globalen Minimum robust wird gegen lokale Minima.

Auch ist es vorteilhaft, wenn ein Vergessensfaktor berücksichtigt wird, um den Einfluß weit zurückliegender schlechter Werte bei der Iteration über alle ermittelten Werte abzuschwächen.

Weiterbildungen des erfindungsgemäßen Verfahrens ergeben sich aus den abhängigen Ansprüchen.

30 Das erfindungsgemäße Verfahren wird anhand der folgenden Figuren weiter erläutert.

Es zeigen

Fig. 1 ein Blockdiagramm, das die einzelnen Komponenten des Verfahrens darstellt,

Fig. 2 eine Skizze, die die Funktion des erfindungsgemäßen Verfahrens veranschaulicht,

Fig. 3 eine Skizze, die die Funktion der Zufallskomponente veranschaulicht,

35 Fig. 4 ein Blockdiagramm, das eine Anwendung des Verfahrens darstellt.

Fig. 1 zeigt das der Erfindung zugrundeliegende Prinzip. Das dargestellte Optimierungsverfahren (Block 1a) erzeugt einen Wert  $x \in R^n$  (x-Wert XW), an dem eine Zielfunktionsauswertung vorgenommen wird (siehe Block 1b). Es ist der Gradient an dieser Stelle  $x$  erforderlich, um einen neuen x-Wert XW zur Funktionsauswertung zu bestimmen. Diese Bestimmung eines x-Wertes XW kann iterativ erfolgen, bis der k-Wert XW mit dem kleinsten Funktionswert gefunden ist. Der Gradient steht i.a. nicht zur Verfügung, so daß er numerisch berechnet werden soll. Bei dem erfindungsgemäßen Verfahren wird der Gradient mittels zurückliegender Funktionsauswertungen rekursiv geschätzt (siehe Block 1c). Dieser Gradient bestimmt die Suchrichtung SR nach dem globalen Minimum in der nächsten Iteration des globalen stochastischen Optimierungsverfahrens (1a).

40 Dabei werden die x-Werte XW nach einem Zufallsprinzip erzeugt, was eine robuste Gradientenschätzung durch den verwendeten RLS-Algorithmus erlaubt. Beispielsweise kann als Eingangssignal für den RLS-Algorithmus weißes Rauschen verwendet werden.

In der Praxis ist die Voraussetzung der zweimal stetigen Differenzierbarkeit der Zielfunktion selten erfüllt. Bei der globalen Minimierung genügt die Kenntnis des Gradienten des zweimal stetig differenzierbaren Anteils, der mit dem RLS-Algorithmus robust bestimmt werden kann.

50 Somit sind Vorteile des erfindungsgemäßen Verfahrens, daß keine zusätzlichen Zielfunktionsauswertungen erforderlich sind und daß die Gradientenschätzung gegen Störungen unempfindlich ist.

An Fig. 2 soll im Folgenden das Verfahren zur globalen Optimierung näher erläutert werden.

Als Voraussetzung sei  $f: R^n \rightarrow R$  eine zweimal stetig differenzierbare Zielfunktion, die ein globales Minimum besitzt. Die Aufgabenstellung besteht in der Bestimmung dieses globalen Minimums.

55 Das erfindungsgemäße Verfahren stellt eine stochastische Methode dar, die auf der numerischen Integration folgender stochastischer Integralgleichung basiert (siehe [2]):

$$60 \quad x_{x_0}^{\varepsilon}(\omega, t) = x_0 - \int_0^t \nabla f(x_{x_0}^{\varepsilon}(\omega, \tau)) d\tau + \varepsilon B(\omega, t) \quad (1),$$

mit  $\omega \in \Omega$

wobei

65  $\Omega$  eine Menge aller Ergebnisse  $\omega$ ,

$t$  einen Zeitpunkt  $t \in R_0^+$ ,

$x_0$  einen Startpunkt  $x_0 \in R^n$ ,

$X$  eine Abb.  $X: \Omega \times R_0^+ \rightarrow R^n$ .

B eine Abb.  $B: \Omega \times \mathbb{R}_0^+ \rightarrow \mathbb{R}^n$   
 $\varepsilon$  einen vorgebbaren Parameter  $\varepsilon > 0$   
 bezeichnen.

Dabei liegt der Wahrscheinlichkeitsraum  $(\Omega, S, P)$  zugrunde und die Abb. B wird als n-dimensionale Brownsche Bewegung bezeichnet. Der Gradient der Zielfunktion wird durch

$$\nabla f(x_{x_0}^\varepsilon(\omega, \tau))$$

und die Zufallssuche durch B dargestellt.

Die stochastische Integralgleichung (1) kombiniert somit die lokale Minimierung (Kurve des steilsten Abstiegs) mit der Zufallssuche (Brownsche Bewegung).

In Fig. 2 wird die numerische Lösung der Gleichung (1) für einen Pfad des stochastischen Prozesses veranschaulicht. Dabei bedeutet der stochastische Prozeß eine Sammlung

$$\{x_{x_0}^\varepsilon(\omega, \bar{t}); \bar{t} \in \mathbb{R}_0^+\}$$

von n-dimensionalen Zufallsvariablen

$$x_{x_0}^\varepsilon: \Omega \rightarrow \mathbb{R}^n; \omega \mapsto x_{x_0}^\varepsilon(\omega, \bar{t})$$

für jedes feste  $\bar{t} \in \mathbb{R}_0^+$

und ein Pfad des stochastischen Prozesses die Funktion

$$x_{x_0}^\varepsilon: [0, \infty[ \rightarrow \mathbb{R}^n, t \mapsto x_{x_0}^\varepsilon(\bar{\omega}, t)$$

für jedes feste  $\bar{\omega} \in \Omega$

Die Zufallssuche ist in Fig. 2 durch einen Kreis ZSK und die Gradientenrichtung durch einen Pfeil GPf symbolisiert. Befindet man sich im Punkt

$$x_{x_0}^\varepsilon(\bar{\omega}, \bar{t}) = x_a,$$

so geht man zur approximativen Berechnung von

$$x_{x_0}^\varepsilon(\bar{\omega}, \bar{t} + T) \text{ mit } T \geq 0 \text{ gemäß } -T \nabla f(x_{x_0}^\varepsilon(\bar{\omega}, \bar{t}))$$

in Richtung des negativen Gradienten, verrauscht dann die Gradientenrichtung gemäß

$$\varepsilon(B(\bar{\omega}, \bar{t} + T) - B(\bar{\omega}, \bar{t})),$$

erhält den neuen Punkt

$$x_{x_0}^\varepsilon(\bar{\omega}, \bar{t} + T) = x_b$$

und setzt die Iteration fort.

Die Wahl des Parameters  $\varepsilon$  wird vorgebbare nach folgenden Kriterien vorgenommen (siehe [2]):

Zeigt sich bei der Betrachtung der bisher berechneten Iterationspunkte, daß bei der numerischen Lösung von (1) die Kurve des steilsten Abstiegs

$$x_{x_0}^0(\bar{\omega}, t) = x_0 - \int_0^t \nabla f(x_{x_0}^0(\bar{\omega}, \tau)) d\tau$$

für ein festes  $\bar{\omega} \in \Omega$  dominiert, so ist  $\varepsilon$  zu erhöhen. Ist hingegen bei der Betrachtung der bisherigen Iterationspunkte die Zufallssuche

$$x_{x_0}^{\varepsilon}(\bar{\omega}, t) \approx x_0 + \varepsilon \cdot B(\bar{\omega}, t)$$

5 für ein festes  $\bar{\omega} \in \Omega$  dominant, so ist  $\varepsilon$  zu verringern.

Somit ist der Parameter  $\varepsilon$  zur Steuerung für eine Regularisierung zwischen lokaler Minimierung und Zufallssuche zu verwenden (Regularisierungsfaktor).

Zur Lösung von Gleichung (1) können numerische Verfahren für Anfangswertprobleme gewöhnlicher Differentialgleichungen eingesetzt werden. Da aber der Algorithmus für Zielfunktionen ausgelegt ist, bei denen weder der exakte Gradient noch die exakte Hessematrix zur Verfügung stehen, wurde als Verfahren zur lokalen Minimierung das explizite Eulerverfahren (Gradientenverfahren) gewählt. Dabei ist wesentlich, daß der benötigte Gradient mittels RLS-Algorithmus geschätzt wird.

Für die Gradientenschätzung wird von folgender Gleichung ausgegangen:

$$15 \quad y = \bar{y}_0 + x' g \quad (2)$$

mit  $y = f(x) \in \mathbb{R}$

wobei

20  $y$  einen Funktionswert von  $x \in \mathbb{R}^n$ ,  $\bar{y}_0$  einen Schnittpunkt der Ebene mit der  $y$ -Achse und  $g$  die zu schätzende Suchrichtung bezeichnen.

Im Folgenden werden  $k+1$  Punkte  $y_0, y_1, y_2, \dots, y_k$  betrachtet. Diese  $k+1$  Punkte können beispielsweise durch Iteration bei einer numerischen Optimierungsaufgabe erzeugt worden sein. Dabei bezeichnen  $k$  die diskrete Zeit,  $x_k$  den aktuellen Punkt und  $x_{k-1}, \dots, x_1, x_0$  weiter zurückliegende Stützpunkte mit  $x_i \in \mathbb{R}^n, i = 1, 2, \dots, k$ . Durch Einsetzen dieser Punkte in (2) ergibt sich das Gleichungssystem

$$30 \quad \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_k \end{pmatrix} = \begin{pmatrix} 1 & x'_0 \\ 1 & x'_1 \\ \vdots & \vdots \\ 1 & x'_k \end{pmatrix} \begin{pmatrix} \bar{y}_{0k} \\ g_k \end{pmatrix} \quad (3),$$

35 wobei  $g_k \in \mathbb{R}^n$  die anhand der genannten  $k+1$  Stützstellen zu schätzende Suchrichtung und  $\bar{y}_{0k} \in \mathbb{R}$  den dazugehörigen Schnittpunkt der Ebene mit der  $y$ -Achse bezeichnen. Eine Subtraktion der einzelnen Gleichungen in (3) bewirkt eine Elimination der Konstanten  $\bar{y}_{0k}$ , die für die Gradientenschätzung nicht von Interesse ist.

$$40 \quad \underbrace{\begin{pmatrix} y_1 - y_0 \\ y_2 - y_1 \\ \vdots \\ y_k - y_{k-1} \end{pmatrix}}_{d_k} = \underbrace{\begin{pmatrix} x'_1 - x'_0 \\ x'_2 - x'_1 \\ \vdots \\ x'_k - x'_{k-1} \end{pmatrix}}_{A_k} g_k \quad (4).$$

Der Gradient  $g_k$  stellt dabei den zu schätzenden Gradienten  $\nabla f(x_k)$  der Zielfunktion  $f$  an der Stelle  $x_k$  dar. Zum Zeitpunkt  $k+1$  lautet die Gleichung (4) dann:

$$50 \quad \underbrace{\begin{pmatrix} y_1 - y_0 \\ y_2 - y_1 \\ \vdots \\ y_{k+1} - y_k \end{pmatrix}}_{d_{k+1}} = \underbrace{\begin{pmatrix} x'_1 - x'_0 \\ x'_2 - x'_1 \\ \vdots \\ x'_{k+1} - x'_k \end{pmatrix}}_{A_{k+1}} g_{k+1} \quad \text{mit } g_{k+1} \in \mathbb{R}^n \quad (4a)$$

60 Das Gleichungssystem (4a) enthält i.a. mehr Stützpunkte (Gleichungen) als Parameter, also gilt  $k+1 > n$ . Wie ein solches überbestimmtes Gleichungssystem in dem erfindungsgemäßen Verfahren gelöst wird, wird im folgenden Abschnitt erläutert.

Der LS-Algorithmus (Least-Square-Algorithmus) wird dazu verwendet, überbestimmte lineare Gleichungssysteme veränderlicher Größe zu lösen.

65 Für ein festes  $k$  aus der Menge der natürlichen Zahlen  $N$  betrachtet man das Gleichungssystem

$$M_k p_k = b_k$$

(5)

mit

$$M_k = \begin{pmatrix} m'_1 \\ m'_2 \\ \vdots \\ m'_k \end{pmatrix} \in \mathbb{R}^{k,n}, \quad m_i \in \mathbb{R}^n, \quad i=1,2,\dots,k,$$

$$p_k \in \mathbb{R}^n \text{ und}$$

$$b_k = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{pmatrix} \in \mathbb{R}^k, \quad \beta_i \in \mathbb{R}, \quad k > n.$$

In der vorliegenden Darstellung werden alle Vektoren als Spaltenvektoren dargestellt.  
I.d.R. existiert kein  $p_k$ , das die Gleichung (5) exakt erfüllt. Das Residuum

$$r(p_k) = b_k - M_k p_k$$

ist also für kein  $p_k$  zum Verschwinden zu bringen. Als Ersatz für die exakte Lösung wird ein  $p_k$  gesucht, das das Residuum möglichst klein werden läßt, z. B. im Sinne der euklidischen Norm soll ein  $\hat{p}_k$  so bestimmt werden, daß gilt

$$\|r(\hat{p}_k)\|_2 \leq \|r(p_k)\|_2 \quad \forall p_k \in \mathbb{R}^n.$$

Dies wird als das lineare Ausgleichsproblem oder die Ausgleichung der Widersprüche nach der Methode der kleinsten Quadrate bezeichnet.

Weiterhin kann der Einfluß weiter zurückliegender Daten verringert werden, damit der zu schätzende Parametervektor  $p_k$  den vorgegebenen Parametern des Gleichungssystems folgen kann. Dieses "Vergessen" bewirkt, daß die Daten mit einem um so kleineren Gewicht versehen werden, je weiter sie zurückliegen.

Ein Möglichkeit hierfür ist das multiplikative Vergessen. Mittels der Gewichtungsmatrix

$$D_k = \begin{pmatrix} d_1 & & 0 \\ & d_2 & \\ & & \ddots \\ 0 & & & d_k \end{pmatrix} \quad \text{mit } d_k = 1, d_{k-v} = \alpha^v, v = 1, 2, \dots, k-1$$

wird die mittlere quadratische Fehlerfunktion gebildet:

$$E_k(p_k) := \frac{1}{2} (b_k - M_k p_k)' D_k (b_k - M_k p_k) \quad (6),$$

wobei  $\alpha$  ein fester vorgebbare Parameter im Intervall  $0 < \alpha \leq 1$  ist und dazu verwendet wird, die einzelnen Summanden in  $E_k(p_k)$  unterschiedlich zu bewerten, und zwar um so schwächer, je weiter ihr Argument  $k-v$  gegenüber  $k$  zurückliegt. Mit dem Parameter  $\alpha$  kann also erreicht werden, daß die zurückliegenden Summanden mit zunehmender Zeit  $k$  immer mehr "in Vergessenheit" geraten. Dieser Effekt tritt nicht auf, wenn  $\alpha = 1$ , d. h.  $D_k = I_k \forall k \in \mathbb{N}$  gewählt wird; dies stellt den Spezialfall der ungewichteten quadratischen Fehlerfunktion dar.

Es soll nun  $p_k$  so bestimmt werden, daß  $E_k(p_k)$  minimal wird. Dazu wird der Gradient der Fehlerfunktion gleich Null gesetzt:

$$\nabla E_k(p_k) = 0 \leftrightarrow M_k' D_k M_k p_k - M_k' D_k b_k = 0.$$

Somit gilt, falls  $M_k' D_k M_k$  invertierbar ist:

$$p_k = R_k^{-1} r_k \quad (7)$$

mit

$$R_k = M_k' D_k M_k \quad (8),$$

$$r_k = M_k^T D_k b_k \quad (9).$$

Setzt man in den Gleichungen (5), (7), (8) und (9)  $M_k = A_k$ ,  $p_k = g_k$  und  $b_k = d_k$ , so ist das Problem der Gradientenschätzung aus Gleichung (4) prinzipiell gelöst. Allerdings ist die durch die Gleichungen (7), (8) und (9) beschriebene Berechnung noch nicht brauchbar, da bei Hinzunahme eines weiteren Vektors  $m_{k+1}$  und eines weiteren Skalars  $\beta_{k+1}$  erneut eine Matrixinversion mit Berechnung von  $R_k^{-1}$  anfällt. Die Größe des Gleichungssystems wächst mit steigendem  $k$  an. Der nachfolgende Algorithmus umgeht diese Schwierigkeit mittels rekursiver Berechnung.

Bei dem RLS-Algorithmus wird das um eine weitere Gleichung ergänzte lineare Gleichungssystem

$$M_{k+1} p_{k+1} = b_{k+1}$$

mit

$$M_{k+1} = \begin{pmatrix} M_k \\ m_{k+1}^T \end{pmatrix}, \quad M_k \in \mathbb{R}^{k,n}, \quad m_{k+1} \in \mathbb{R}^n \quad \text{und}$$

$$b_{k+1} = \begin{pmatrix} b_k \\ \beta_{k+1} \end{pmatrix}, \quad b_k \in \mathbb{R}^n, \quad \beta_{k+1} \in \mathbb{R}$$

betrachtet. Dabei stellt  $p_{k+1}$  den neuen zur Zeit  $k+1$  gesuchten Vektor dar, mit dem jetzt  $k+1$  Bedingungen bestmöglich zu erfüllen sind. Die Zeilenzahl der Matrix  $M_{k+1}$  und die Länge des Vektors  $b_{k+1}$  sind also gegenüber  $M_k$  und  $b_k$  um "1" angewachsen.

Das wesentliche Merkmal des RLS-Algorithmus ist nun die rekursive Ermittlung der inversen Matrix  $R_k^{-1}$ . Daher ist das Ziel der nachfolgenden Betrachtungen, eine Lösung folgender Form zu finden:

$$\begin{pmatrix} \text{neuer} \\ \text{Schätzwert} \end{pmatrix} = \begin{pmatrix} \text{alter} \\ \text{Schätzwert} \end{pmatrix} + \begin{pmatrix} \text{Korrektur} - \\ \text{vektor} \end{pmatrix} \cdot \begin{pmatrix} \text{Korrektur} - \\ \text{faktor} \end{pmatrix}$$

In den folgenden Schritten wird das Aufstellen der rekursiven Gleichungen näher erläutert. Aus den Gleichungen (7), (8) und (9) folgt für den Zeitpunkt  $k+1$

$$p_{k+1} = R_{k+1}^{-1} \cdot \eta_{k+1} \quad (10),$$

$$R_{k+1} = M_{k+1}^T D_{k+1} M_{k+1} \quad (11),$$

$$\eta_{k+1} = M_{k+1}^T D_{k+1} b_{k+1} \quad (12).$$

Durch Ausmultiplizieren ergeben sich aus (11) und (12) folgende Gleichungen:



$$R_{k+1} = \sum_{v=1}^{k+1} \alpha^{k+1-v} m_v m_v'$$

mit  $m_v' \in \mathbb{R}$  ( $v$ -te Zeile von  $M_{k+1}$ )

$$\Leftrightarrow R_{k+1} = \alpha \sum_{v=1}^k \alpha^{k-v} m_v m_v' + m_{k+1} m_{k+1}'$$

$$\Leftrightarrow R_{k+1} = \alpha R_k + m_{k+1} m_{k+1}' \quad (13)$$

und

$$\eta_{k+1} = \sum_{v=1}^{k+1} \alpha^{k+1-v} \beta_v m_v$$

mit  $\beta_v \in \mathbb{R}^n$  ( $v$ -te Komponente von  $b_{k+1}$ )

$$\Leftrightarrow \eta_{k+1} = \alpha \sum_{v=1}^k \alpha^{k-v} \beta_v m_v + \beta_{k+1} m_{k+1}$$

$$\Leftrightarrow \eta_{k+1} = \alpha \eta_k + \beta_{k+1} m_{k+1} \quad (14)$$

Gleichung (13) stellt eine Rekursion für die Berechnung der Matrix  $R_{k+1}$  dar. Daraus läßt sich unmittelbar eine Rekursion für die inverse Matrix  $R_{k+1}^{-1}$  bestimmen (siehe [1]):

$$R_{k+1}^{-1} = \frac{1}{\alpha} \left( R_k^{-1} - \frac{R_k^{-1} \cdot m_{k+1} \cdot m_{k+1}' \cdot R_k^{-1}}{\alpha + m_{k+1}' \cdot R_k^{-1} \cdot m_{k+1}} \right) \quad (15)$$

d. h. man kann  $R_{k+1}^{-1}$  ohne Matrixinversion aus der inversen Matrix  $R_k^{-1}$  des vorherigen Schrittes berechnen. Mit der Abkürzungen

$$h_{k+1} = \frac{R_k^{-1} \cdot m_{k+1}}{\alpha + m_{k+1}' \cdot R_k^{-1} \cdot m_{k+1}}$$

erhält man

$$R_{k+1}^{-1} = \frac{1}{\alpha} \left( R_k^{-1} - h_{k+1} m_{k+1}' R_k^{-1} \right) \quad (16)$$

Für den zu schätzenden Parametervektor ergibt sich durch Einsetzen von (14) und (16) in (10) und nach weiteren Umformungen die Gleichung (siehe auch [1])

$$p_{k+1} = p_k + h_{k+1} (\beta_{k+1} - m_{k+1}' p_k)$$

$$\text{bzw. mit } e_k = \beta_{k+1} - m_{k+1}' p_k$$

$$p_{k+1} = p_k + h_{k+1} e_k \quad (17)$$

Für das rekursive Verfahren werden vorteilhaft die Startwerte  $p_1$  und  $R_1^{-1}$  benötigt. Als Startwerte werden diejenigen Werte verwendet, die einen möglichst geringen Einfluß auf das Schätzergebnis nehmen: Es empfiehlt sich als Startwerte  $p_1 \approx 0$  und  $R_1^{-1} = \eta I_n$  mit  $\eta \gg 1$  zu verwenden. Die dadurch entstehende Ungenauigkeit beeinflußt den Vektor  $p_k$ . In der Praxis ist es vorteilhaft, den Parameter  $\eta$  möglichst groß zu wählen.

Der RLS-Algorithmus gewinnt seine Information über die Datenmatrix  $M_k$  und den Datenvektor  $b_k$  und speichert diese Information in der Matrix  $R_k^{-1}$ . Führt man eine Eigenwert-Eigenvektor-Zerlegung der Matrix  $R_k$  durch, so ergibt

sich

$$R_k = T_k \Lambda_k T_k^T \quad (18)$$

5 mit  $\Lambda_k = \text{diag}(\lambda_{1k}, \lambda_{2k}, \dots, \lambda_{n_k})$ ,

wobei  $\lambda_{ik}; i = 1, 2, \dots, n$  die Eigenwerte der Matrix  $R_k$  darstellen und die Matrix

$$10 \quad T_k = (u_{1k}, u_{2k}, \dots, u_{n_k}) \quad (19)$$

aus den Eigenvektoren  $u_{ik}; i = 1, 2, \dots, n$  der Matrix  $R_k$  besteht.

Aus den Gleichungen (13) und (18) folgt für den Fall des multiplikativen Vergessens

$$15 \quad R_{k+1} = T_k \begin{pmatrix} \alpha \lambda_{1k} & & & \\ & \alpha \lambda_{2k} & & \\ & & \ddots & \\ & & & \alpha \lambda_{n_k} \end{pmatrix} T_k^T + m_{k+1} m_k^T + 1$$

mit

$$25 \quad \tilde{R}_k = T_k \begin{pmatrix} \alpha \lambda_{1k} & & & \\ & \alpha \lambda_{2k} & & \\ & & \ddots & \\ & & & \alpha \lambda_{n_k} \end{pmatrix} T_k^T \quad (20)$$

Für multiplikatives Vergessen gilt also folgende Beziehung für die Eigenwerte  $\lambda_{ik}$  von  $R_k$  und die Eigenwerte  $\tilde{\lambda}_{ik}$  von  $\tilde{R}_k$ :  $\tilde{\lambda}_{ik} = \alpha R_k$

$$35 \quad \tilde{\lambda}_{ik} = \alpha \cdot \lambda_{ik} \quad \text{mit } i=1, 2, \dots, n.$$

Da die Matrix  $R_k$  per Definition symmetrisch und positiv (semi-)definit ist, folgt

$$40 \quad \lambda_{ik} \geq 0 \text{ für } i = 1, 2, \dots, n$$

und

$$T_k^T T_k = I_n.$$

45 Durch Inversion von  $\tilde{R}_k$  erhält man

$$50 \quad \tilde{R}_k^{-1} = T_k \begin{pmatrix} \frac{1}{\alpha \lambda_{1k}} & & & \\ & \frac{1}{\alpha \lambda_{2k}} & & \\ & & \ddots & \\ & & & \frac{1}{\alpha \lambda_{n_k}} \end{pmatrix} T_k^T \quad (21)$$

60 bzw. für die Eigenwerte

$$\frac{1}{\tilde{\lambda}_{ik}} = \frac{1}{\alpha \lambda_{ik}} \quad \text{für } i=1, 2, \dots, n \quad (22)$$

65 Es ist vorteilhaft, wenn anstelle der Gleichung (22) folgende Gleichung verwendet wird:

$$\frac{1}{\tilde{\lambda}_{i_k}} = \frac{1}{\lambda_{i_k}} + \frac{1-\alpha}{\alpha} \left( \frac{1}{\lambda_{i_k}} \right)^{1-\delta} \quad \text{für } i=1,2,\dots,n \quad (23),$$

wobei je nach Wahl von  $\delta$  verschiedene Einstellungen für das Vergessen möglich sind:

- multiplikatives Vergessen wie in Gleichung (22) für  $\delta = 0$ ,

• additives Vergessen mit  $\delta=1$ :  $\frac{1}{\tilde{\lambda}_{i_k}} = \frac{1}{\lambda_{i_k}} + \frac{1-\alpha}{\alpha}$

Durch Einsetzen der Diagonalelemente in Gleichung (20) mit den Kehrwerten der Eigenwerte gemäß Gleichung (23) ergibt sich

$$\tilde{R}_k = T_k \begin{pmatrix} \frac{\lambda_{1_k}}{1 + \frac{1-\alpha}{\alpha} \lambda_{1_k}^\delta} & & & \\ & \frac{\lambda_{2_k}}{1 + \frac{1-\alpha}{\alpha} \lambda_{2_k}^\delta} & & \\ & & \ddots & \\ & & & \frac{\lambda_{n_k}}{1 + \frac{1-\alpha}{\alpha} \lambda_{n_k}^\delta} \end{pmatrix} T_k \quad (24).$$

Durch Ersetzen von  $\tilde{R}_k$  anstelle von  $\alpha R_k$  in (13) ergibt sich

$$R_{k+1} = \tilde{R}_k + m_{k+1} m'_{k+1}.$$

Ersetzt man die Diagonalelemente der Matrix aus (21) durch die Eigenwerte der Gleichung (23), so ergibt sich

$$\tilde{R}_k^{-1} = T_k \begin{pmatrix} \frac{1}{\lambda_{1_k}} + \frac{1-\alpha}{\alpha} \left( \frac{1}{\lambda_{1_k}} \right)^{1-\delta} & & & \\ & \ddots & & \\ & & \ddots & \\ & & & \frac{1}{\lambda_{n_k}} + \frac{1-\alpha}{\alpha} \left( \frac{1}{\lambda_{n_k}} \right)^{1-\delta} \end{pmatrix} T_k \quad (25).$$

Aus [1] folgt eine neue rekursive Beziehung

$$R_{k+1}^{-1} = \tilde{R}_k^{-1} - \frac{\tilde{R}_k^{-1} \cdot m_{k+1} \cdot m'_{k+1} \cdot \tilde{R}_k^{-1}}{1 + m'_{k+1} \cdot \tilde{R}_k^{-1} \cdot m_{k+1}} \quad (26).$$

Mit der Abkürzung

$$h_{k+1} = \frac{\tilde{R}_k^{-1} \cdot m_{k+1}}{1 + m'_{k+1} \cdot \tilde{R}_k^{-1} \cdot m_{k+1}} \quad (27)$$

wird die Rekursion berechnet:

$$R_{k+1}^{-1} = \tilde{R}_k^{-1} - h_{k+1} \cdot m'_{k+1} \cdot \tilde{R}_k^{-1} \quad (28).$$

Zur Berechnung der gesuchten Größe  $p_{k+1}$  setzt man wie beim multiplikativen Vergessen Gleichung (17) mit der Abkürzung

$$e_k = \beta_{k+1} - m'_{k+1} p_k \quad (29)$$

zu

$$p_{k+1} = p_k + h_{k+1} e_k \quad (30).$$

Nachfolgend wird nochmals zusammengefaßt, wie mittels RLS-Algorithmus der Gradient einer Zielfunktion  $f$  an einer bestimmten Stelle  $x_i$  bestimmt wird.

Die Iteration startet im Schritt  $k = 1$ . Es werden die Eigenwerte und Eigenvektoren der inversen Matrix  $R_k^{-1}$  bestimmt und die Transformationsmatrix  $T_k$  aufgestellt. Dann wird nach Gleichung (25) die inverse Matrix  $\tilde{R}_k^{-1}$  berechnet. Die Parameter  $\alpha$  und  $\delta$  sind vorgebar gemäß Gleichung (23). Daraufhin sind die Abkürzungen  $e_k$  und  $h_{k+1}$  zu berechnen nach den Gleichungen (29) und (27). Dabei sind  $m'_{k+1}$  durch  $x'_{k+1} - x'_k$  und  $\beta_{k+1}$  durch  $y_{k+1} - y_k$  gemäß Gleichung (4a) zu ersetzen. Danach wird  $p_{k+1}$  laut Gleichung (30) mittels  $p_k$  und den Abkürzungen  $e_k$  und  $h_{k+1}$  ausgewertet. Schließlich wird die inverse Matrix  $R_{k+1}^{-1}$  gemäß Gleichung (28) bestimmt und gegebenenfalls mit dem nächsten Iterationsschritt fortgefahren.

Als Ergebnis stellt  $p_{k+1}$  den gesuchten Gradienten  $g_{k+1}$  zur Zeit  $k+1$  gemäß Gleichung (4a) dar.

In Fig. 3 ist eine Skizze dargestellt, die das erfindungsgemäße Verfahren anhand einer Zielfunktion  $f(x)$  mit lokalen Minima zeigt. Durch die Kombination mit der Zufallssuche endet das Verfahren nicht in einem lokalen Minimum LM, sondern springt unter Berücksichtigung der Brownschen Bewegung und des negativen Gradienten aus dem lokalen Minimum LM in Richtung globales Minimum GM. Das erfindungsgemäße Verfahren durchläuft also im steilen Bereich die lokalen Minima und bleibt im flachen Bereich über mehrere Iterationen hinweg in dem globalen Minimum. Der Pfad PF deutet die Suchrichtung des erfindungsgemäßen Verfahrens an.

Fig. 4 zeigt ein Blockdiagramm, das ein Anwendungsbeispiel für eine globale Optimierung mit Gradientenschätzung darstellt. Als Anwendungsbeispiel dient ein Kühlstreckenmodell KSM der Stahlindustrie. Auf der Basis von Meßdaten MD werden Modellparameter optimiert, d. h. der Modellfehler MF unter Berücksichtigung der Gesamtheit aller Meßdaten MD minimiert.

Dazu werden die Meßdaten MD in das Kühlstreckenmodell KSM eingegeben. Die vom Kühlstreckenmodell KSM gelieferten Ergebnisse werden mit den tatsächlichen Ergebnissen der Kühlstrecke GT verglichen und in Form eines Modellfehlers MF der globalen Optimierung mit Gradientenschätzung GLOPGS zugeführt. Hier werden die Parameter des Kühlstreckenmodells KSM zur Minimierung des Modellfehlers MF optimiert.

Im Rahmen dieses Dokuments wurden folgende Veröffentlichungen zitiert:

- [1] Unbehauen, R.: Systemtheorie. R. Oldenbourg Verlag, Wien 1993. S. 537-545.
- [2] Schäffler, S.: Unconstructed Global Optimization Using Stochastic Integral Equations, Optimization 1995, Vol.35, S. 43-55.

#### Patentansprüche

1. Verfahren zur Optimierung von Prozeßparametern auf der Basis einer Gradientenschätzung auf einem Rechner,
  - a) bei dem in einem Schritt  $k+1$  ein neuer  $x$ -Wert  $x_{k+1}$  in Abhängigkeit von einer Suchrichtung  $-g_k$  berechnet wird gemäß

$$x_{k+1} = x_k - g_k f(x_k),$$

wobei

$f(x_k)$  den zugehörigen Funktionswert aus einer vorgegebenen Funktion  $f$  bezeichnet,

- b) bei dem ein neuer Funktionswert  $y_{k+1}$  berechnet wird gemäß

$$y_{k+1} = f(x_{k+1}),$$

- c) bei dem ein neuer Gradient  $g_{k+1}$  basierend auf einem RLS-Algorithmus berechnet wird,
- d) bei dem der neue  $x$ -Wert nach einem Zufallsprinzip erzeugt wird (Rauschen).

2. Verfahren nach Anspruch 1, bei dem die Schritte 1a) bis 1c) zu mindestens einem nächsten diskreten Zeitpunkt  $k$  iteriert werden.

3. Verfahren nach einem der obigen Ansprüche, bei dem bei der Berechnung des Gradienten ein vorgegebener Vergessensfaktor  $\alpha$  verwendet wird, der  $x$ -Werte zu verschiedenen Zeitpunkten  $k$  unterschiedlich bewertet.

Hierzu 3 Seite(n) Zeichnungen

- Leerseite -

FIG 1

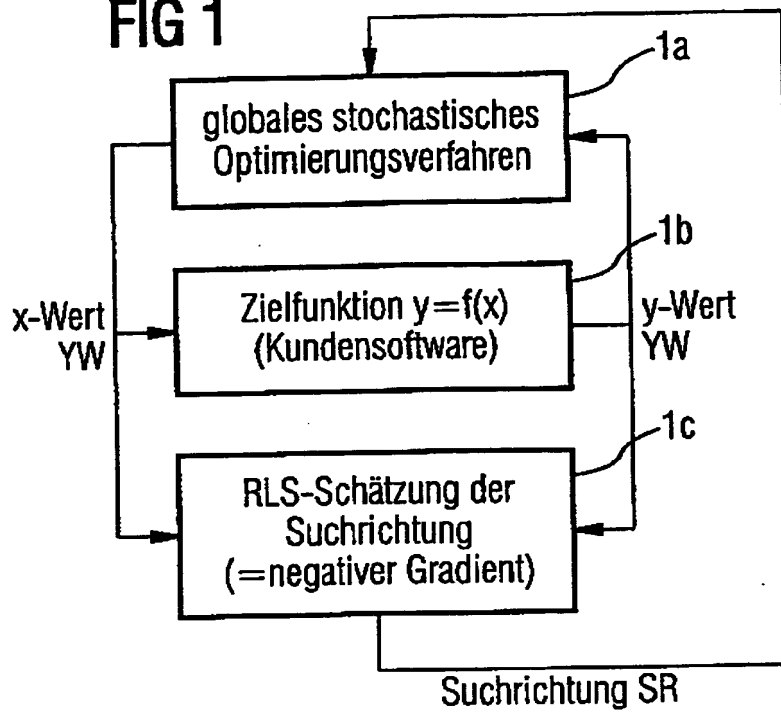


FIG 4

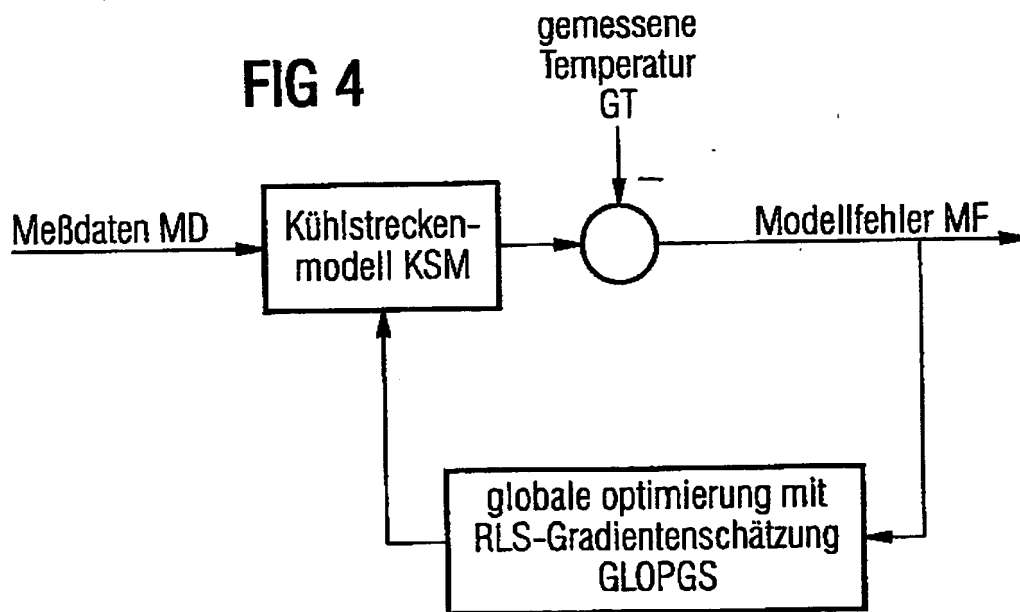


FIG 2

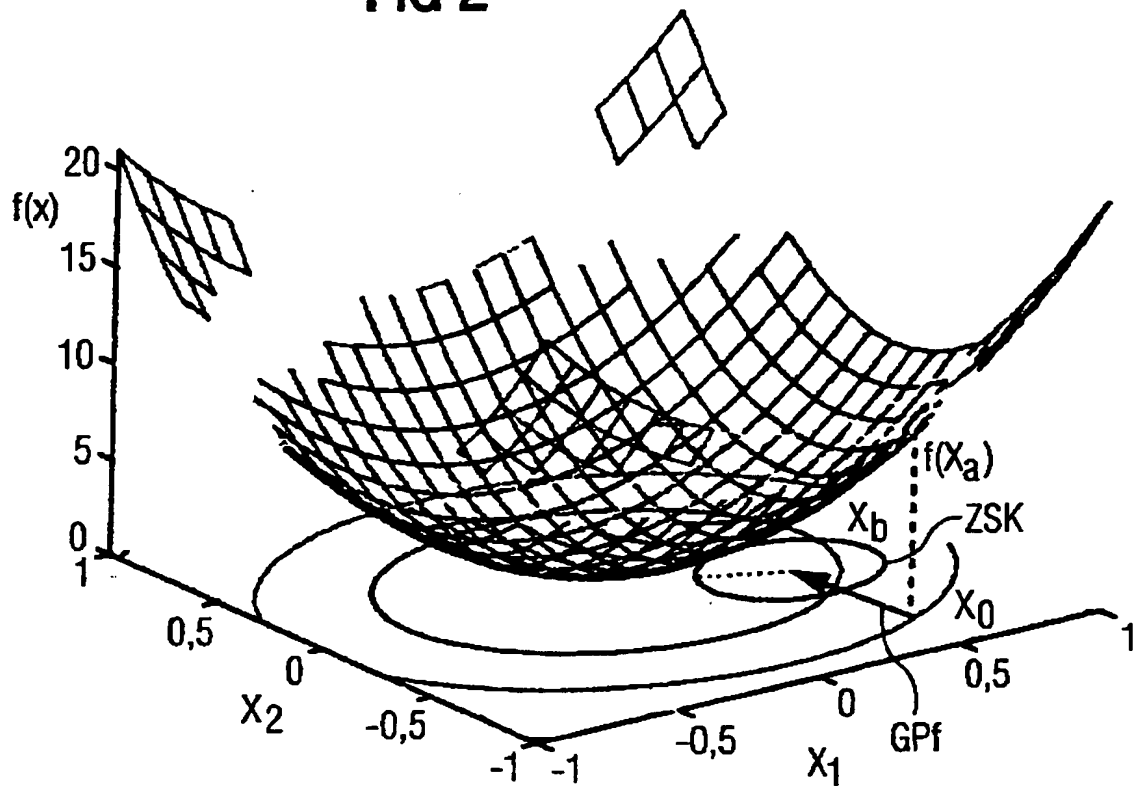


FIG 3

